2.1 Discrete Random Variables and their Probability Mass Functions

Often one is interested in a function of the outcome of an experiment. For instance, the difference of the roll of two (distinguished) dice is a function of the outcome of an experiment. A function of the outcome of an experiment is a "random variable".

Def A random variable Y is a function from a sample space to the real numbers:

$$Y: S \longrightarrow \mathbb{R}.$$

Rem A random variable is not random and not a variable! It is only "random" in the sense that its value at a sample point is the value on the outcome of an experiment, which in turn cannot be predicted.

<u>Def</u> A random variable is <u>discrete</u> if it takes on at most countably many values. ("At most countably many" means at most as many as the set of natural numbers 1,2,3,..., the range is <u>not</u> required to be the set of natural numbers.)

<u>Def</u> Let Y be a random variable, and let y be a real number. The <u>event that Y takes on the value y</u> is the set

$$(Y=y) := \{s \in S \mid Y(s) = y\}$$

= the <u>Y-preimage of y</u>

$= Y^{-1}(y)$ in math notation

(does <u>not</u> mean $\frac{1}{Y(y)}$, $Y(y)$ is not even defined, also does <u>not</u> mean Y is an invertible function, it does <u>not</u> mean the inverse function).

<u>Def</u> The <u>probability mass function</u> <u>of a discrete random variable</u> $Y: S \longrightarrow \mathbb{R}$ is the function $f: \text{image}(Y) \longrightarrow [0,1]$ defined by

$$f(y) := P(Y=y).$$

The probability mass function (<u>p.m.f.</u>) of Y is sometimes referred to as <u>the probability distribution of Y</u>, although that term is ambiguous.

<u>Rem</u> We can also think of the p.m.f. of a discrete random variable as a function

$$f: \mathbb{R} \longrightarrow [0,1]$$

(Notice: domain here is all reals, rather than just image(Y).)

by defining f to be zero outside of image(Y). That is

$$f(y) := \begin{cases} P(Y=y) & \text{if } y \in \text{image}(Y) \\ 0 & \text{otherwise.} \end{cases}$$

<u>Example</u> (Of a Discrete Random Variable, its p.m.f., and Visualizations of its p.m.f.)

Consider the roll of two distinguished 3-sided dice, a red one and a blue one. The sample space is $S = \{(1,1), (1,2), (1,3),$
$(2,1), (2,2), (2,3),$
$(3,1), (3,2), (3,3)\}$.

Let $Y : S \to \mathbb{R}$ be the random variable

$$Y(a,b) := a - b.$$

Let's fully compute Y.

$Y(1,1) = 1-1 = 0$

$Y(1,2) = 1-2 = -1$

$Y(1,3) = 1-3 = -2$

$Y(2,1) = 2-1 = 1$

$Y(2,2) = 2-2 = 0$

$Y(2,3) = 2-3 = -1$

$Y(3,1) = 3-1 = 2$

$Y(3,2) = 3-2 = 1$

$Y(3,3) = 3-3 = 0$

$\Longrightarrow$ Y is a <u>discrete</u> random variable because it takes on at most countably many values, in fact only finitely many.